Ref #	Hits	Search Query	DBs	Default Operator	Plurals	Time Stamp
L2	6	"text summary" & "occurrence frequency"	US-PGPUB; USPAT	OR	ON	2006/01/04 12:24
L3	29	"text summary" & (vector same space)	US-PGPUB; USPAT	OR	ON	2006/01/04 13:56
L4	0	"text summary" & (sentece same extract$3)	US-PGPUB; USPAT	OR	ON	2006/01/04 13:56
L5	49	"text summary" & (sentence same extract$3)	US-PGPUB; USPAT	OR	ON	2006/01/04 13:57
L6	44	"text summary" & (sentence near3 extract$3)	US-PGPUB; USPAT	OR	ON	2006/01/04 13:56
L7	1	"text summary" & (sentence same extract$3 same weight$3 same frequency)	US-PGPUB; USPAT	OR	ON	2006/01/04 14:01
L8	4	"text summary" & SVD	US-PGPUB; USPAT	OR	ON	2006/01/04 14:00
L9	1087	frequency & SVD	US-PGPUB; USPAT	OR	ON	2006/01/04 14:00
L10	205	frequency same SVD	US-PGPUB; USPAT	OR	ON	2006/01/04 14:00
L11	2	frequency same SVD same sentence	US-PGPUB; USPAT	OR	ON	2006/01/04 14:01
L12	26	9 & (document near3 summar$4)	US-PGPUB; USPAT	OR	ON	2006/01/04 14:02
L13	6	"text summary" & "occurrence frequency"	US-PGPUB; USPAT	OR	ON	2006/01/04 14:06
L14	29	"text summary" & (vector same space)	US-PGPUB; USPAT	OR	ON	2006/01/04 14:06
L15	0	"text summary" & (sentece same extract$3)	US-PGPUB; USPAT	OR	ON	2006/01/04 14:06
L16	49	"text summary" & (sentence same extract$3)	US-PGPUB; USPAT	OR	ON	2006/01/04 14:06
L17	44	"text summary" & (sentence near3 extract$3)	US-PGPUB; USPAT	OR	ON	2006/01/04 14:06
L18	1	"text summary" & (sentence same extract$3 same weight$3 same frequency)	US-PGPUB; USPAT	OR	ON	2006/01/04 14:06
L19	4	"text summary" & SVD	US-PGPUB; USPAT	OR	ON	2006/01/04 14:06
L20	1087	frequency & SVD	US-PGPUB; USPAT	OR	ON	2006/01/04 14:06
L21	205	frequency same SVD	US-PGPUB; USPAT	OR	ON	2006/01/04 14:06
L22	2	frequency same SVD same sentence	US-PGPUB; USPAT	OR	ON	2006/01/04 14:06

L23	26	9 & (document near3 summar$4)	US-PGPUB; USPAT	OR	ON	2006/01/04 14:06

PORTAL USPTO

Subscribe (Full Service) Register (Limited Service, Free) Login

Search: ◉ The ACM Digital Library ○ The Guide

sentences and term vector frequency SEARCH

THE ACM DIGITAL LIBRARY Advanced Search Search Tips

Enter words, phrases or names below. Surround phrases or full names with double quotation marks.

Search within Results: **44,965** found

sentences and term vector frequency

Clear result set

SEARCH

Desired Results:

must have **all** of the words or phrases

must have **any** of the words or phrases

must have **none** of the words or phrases

Name or Affiliation:

Authored by: ◉ all ○ any ○ none

Edited by: ◉ all ○ any ○ none

Reviewed by: ◉ all ○ any ○ none

Only search in:*

○ Title ○ Abstract ○ Review ◉ All Information

SEARCH

*Searches will be performed on all available information, including full text where available, unless specified above.

ISBN / ISSN: ◉ Exact ○ Expand

DOI: ◉ Exact ○ Expand

SEARCH

Published:

By: ◉ all ○ any ○ none

In: ◉ all ○ any ○ none

Since: Month Year

Before: Month Year

As: Any type of publication

Conference Proceeding:

Sponsored By:

Conference Location:

Conference Year: yyyy

SEARCH

Classification: (CCS) ☐ **Primary Only**

Classified as: ◉ all ○ any ○ none

Results must have accessible:

☐ Full Text ☐ Abstract ☐ Review

Subject Descriptor: ◉ all ○ any ○ none

Keyword Assigned: ◉ all ○ any ○ none

The ACM Portal is published by the Association for Computing Machinery. Copyright © 2006 ACM, Inc.
Terms of Usage Privacy Policy Code of Ethics Contact Us

PORTAL USPTO

Subscribe (Full Service) Register (Limited Service, Free) Login

Search: The ACM Digital Library The Guide

document summary using sentence term vector frequency

THE ACM DIGITAL LIBRARY

Feedback Report a problem Satisfaction survey

Terms used
document summary using sentence term vector frequency

Found **76,076** of **169,166**

Sort results by: relevance
Display results: expanded form
Save results to a Binder
Search Tips
Open results in a new window

Try an Advanced Search
Try this search in The ACM Guide

Results 1 - 20 of 200 Result page: **1** 2 3 4 5 6 7 8 9 10 next
Best 200 shown

Relevance scale

1 Technique for automatically correcting words in text

Karen Kukich
December 1992 **ACM Computing Surveys (CSUR)**, Volume 24 Issue 4
Publisher: ACM Press

Full text available: pdf(6.23 MB) Additional Information: full citation, abstract, references, citings, index terms, review

Research aimed at correcting words in text has focused on three progressively more difficult problems:(1) nonword error detection; (2) isolated-word error correction; and (3) context-dependent work correction. In response to the first problem, efficient pattern-matching and n-gram analysis techniques have been developed for detecting strings that do not appear in a given word list. In response to the second problem, a variety of general and application-specific spelling cor ...

Keywords: n-gram analysis, Optical Character Recognition (OCR), context-dependent spelling correction, grammar checking, natural-language-processing models, neural net classifiers, spell checking, spelling error detection, spelling error patterns, statistical-language models, word recognition and correction

2 Generic text summarization using relevance measure and latent semantic analysis

Yihong Gong, Xin Liu
September 2001 **Proceedings of the 24th annual international ACM SIGIR conference on Research and development in information retrieval**
Publisher: ACM Press

Full text available: pdf(173.90 KB) Additional Information: full citation, abstract, references, citings, index terms

In this paper, we propose two generic text summarization methods that create text summaries by ranking and extracting sentences from the original documents. The first method uses standard IR methods to rank sentence relevances, while the second method uses the latent semantic analysis technique to identify semantically important sentences, for summary creations. Both methods strive to select sentences that are highly ranked and different from each other. This is an attempt to create a summa ...

Keywords: generic text summarization, relevance measure, semantic analysis

3 Summarization: Web-page summarization using clickthrough data

Jian-Tao Sun, Dou Shen, Hua-Jun Zeng, Qiang Yang, Yuchang Lu, Zheng Chen
August 2005 **Proceedings of the 28th annual international ACM SIGIR conference on**

Research and development in information retrieval SIGIR '05

Publisher: ACM Press

Full text available: pdf(157.70 KB) Additional Information: full citation, abstract, references, index terms

Most previous Web-page summarization methods treat a Web page as plain text. However, such methods fail to uncover the full knowledge associated with a Web page needed in building a high-quality summary, because many of these methods do not consider the hidden relationships in the Web. Uncovering the hidden knowledge is important in building good Web-page summarizers. In this paper, we extract the extra knowledge from the clickthrough data of a Web search engine to improve Web-page summarization ...

Keywords: clickthrough data, generic web-page summarization, latent semantic analysis, thematic lexicon

4 Automatic phrase indexing for document retrieval

J. Fagan

November 1987 **Proceedings of the 10th annual international ACM SIGIR conference on Research and development in information retrieval**

Publisher: ACM Press

Full text available: pdf(1.10 MB) Additional Information: full citation, abstract, references, citings, index terms

An automatic phrase indexing method based on the term discrimination model is described, and the results of retrieval experiments on five document collections are presented. Problems related to this non-syntactic phrase construction method are discussed, and some possible solutions are proposed that make use of information about the syntactic structure of document and query texts.

5 Summarization: Topic themes for multi-document summarization

Sanda Harabagiu, Finley Lacatusu

August 2005 **Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval SIGIR '05**

Publisher: ACM Press

Full text available: pdf(245.85 KB) Additional Information: full citation, abstract, references, index terms

The problem of using topic representations for multi-document summarization (MDS) has received considerable attention recently. In this paper, we describe five different topic representations and introduce a novel representation of topics based on topic themes. We present eight different methods of generating MDS and evaluate each of these methods on a large set of topics used in past DUC workshops. Our evaluation results show a significant improvement in the quality of summaries based on topic ...

Keywords: summarization, topic themes

6 Information access and retrieval: Multiple related document summary and navigation using concept hierarchies for mobile clients

D. L. Chan, R. W. P. Luk, W. K. Mak, H. V. Leong, E. K. S. Ho, Q. Lu

March 2002 **Proceedings of the 2002 ACM symposium on Applied computing**

Publisher: ACM Press

Full text available: pdf(660.36 KB) Additional Information: full citation, abstract, references, index terms

Mobile clients have limited display and navigation capabilities. To browse a set of documents, an intuitive method is to navigate through concept hierarchies. To reduce semantic loading for each term that represents the concepts and the cognitive loading of users due to the limited display, similar documents are grouped together before concept hierarchies are constructed for each document group. Since the concept hierarchies only represent the salient concepts in the documents, term extraction i ...

Keywords: browsing, concept hierarchy, information access, mobile agent, mobile computing, navigation, summarization

7 Entity-based cross-document coreferencing using the Vector Space Model
Amit Bagga, Breck Baldwin
August 1998 **Proceedings of the 17th international conference on Computational linguistics - Volume 1 , Proceedings of the 36th annual meeting on Association for Computational Linguistics - Volume 1**
Publisher: Association for Computational Linguistics , Association for Computational Linguistics
Full text available: pdf(593.52 KB) Publisher Site
Additional Information: full citation, abstract, references, citings

Cross-document coreference occurs when the same person, place, event, or concept is discussed in more than one text source. Computer recognition of this phenomenon is important because it helps break "the document boundary" by allowing a user to examine information about a particular entity from multiple text sources at the same time. In this paper we describe a cross-document coreference resolution algorithm which uses the Vector Space Model to resolve ambiguities between people having the same ...

8 Text classification: Web-page classification through summarization
Dou Shen, Zheng Chen, Qiang Yang, Hua-Jun Zeng, Benyu Zhang, Yuchang Lu, Wei-Ying Ma
July 2004 **Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval SIGIR '04**
Publisher: ACM Press
Full text available: pdf(225.80 KB) Additional Information: full citation, abstract, references, index terms, review

Web-page classification is much more difficult than pure-text classification due to a large variety of noisy information embedded in Web pages. In this paper, we propose a new Web-page classification algorithm based on Web summarization for improving the accuracy. We first give empirical evidence that ideal Web-page summaries generated by human editors can indeed improve the performance of Web-page classification algorithms. We then propose a new Web summarization-based classification algorithm ...

Keywords: content body, web page categorization, web page summarization

9 Summarizing text documents: sentence selection and evaluation metrics
Jade Goldstein, Mark Kantrowitz, Vibhu Mittal, Jaime Carbonell
August 1999 **Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval**
Publisher: ACM Press
Full text available: pdf(273.01 KB) Additional Information: full citation, references, citings, index terms

10 Set-based vector model: An efficient approach for correlation-based ranking
Bruno Pôssas, Nivio Ziviani, Wagner Meira, Berthier Ribeiro-Neto
October 2005 **ACM Transactions on Information Systems (TOIS)**, Volume 23 Issue 4
Publisher: ACM Press
Full text available: pdf(800.89 KB) Additional Information: full citation, abstract, references, index terms

This work presents a new approach for ranking documents in the vector space model. The novelty lies in two fronts. First, patterns of term co-occurrence are taken into account and are processed efficiently. Second, term weights are generated using a data mining technique called association rules. This leads to a new ranking mechanism called the *set-based vector model*. The components of our model are no longer index terms but index termsets, where a termset is a set of index terms. Termset ...

Keywords: Information retrieval models, association rule mining, correlation-based

ranking, data mining, weighting index term co-occurrences

11 Using syntactic analysis in a document retrieval system that uses signature files

R. Sacks-Davis, P. Wallis, R. Wilkinson

December 1989 **Proceedings of the 13th annual international ACM SIGIR conference on Research and development in information retrieval**

Publisher: ACM Press

Full text available: pdf(1.15 MB) Additional Information: full citation, abstract, references, citings, index terms

Our work involves the study of the extent to which natural language processing techniques aid the automatic indexing and retrieval of documents. In this paper we describe the use of signature files in large text retrieval systems. We show that good performance can be obtained without requiring the significant overheads required for the inverted file technique. We examine the use of syntactic analysis of the text in all stages of retrieval and argue that an initial Boolean query should be pe ...

12 Multidocument summarization: An added value to clustering in interactive retrieval

Manuel J. Maña-López, Manuel De Buenaga, José M. Gómez-Hidalgo

April 2004 **ACM Transactions on Information Systems (TOIS)**, Volume 22 Issue 2

Publisher: ACM Press

Full text available: pdf(199.91 KB) Additional Information: full citation, abstract, references, index terms, review

A more and more generalized problem in effective information access is the presence in the same corpus of multiple documents that contain similar information. Generally, users may be interested in locating, for a topic addressed by a group of similar documents, one or several particular aspects. This kind of task, called instance or aspectual retrieval, has been explored in several TREC Interactive Tracks. In this article, we propose in addition to the classification capacity of clustering techn ...

Keywords: Multidocument summarization, topic segmentation

13 Extracting sentence segments for text summarization: a machine learning approach

Wesley T. Chuang, Jihoon Yang

July 2000 **Proceedings of the 23rd annual international ACM SIGIR conference on Research and development in information retrieval**

Publisher: ACM Press

Full text available: pdf(945.26 KB) Additional Information: full citation, abstract, references, citings, index terms

With the proliferation of the Internet and the huge amount of data it transfers, text summarization is becoming more important. We present an approach to the design of an automatic text summarizer that generates a summary by extracting sentence segments. First, sentences are broken into segments by special cue markers. Each segment is represented by a set of predefined features (e.g. location of the segment, average term frequencies of the words occurring in the segment, number of title words ...

Keywords: machine learning, sentence segment extraction, text summarization

14 Query expansion using lexical-semantic relations

Ellen M. Voorhees

August 1994 **Proceedings of the 17th annual international ACM SIGIR conference on Research and development in information retrieval**

Publisher: Springer-Verlag New York, Inc.

Full text available: pdf(692.34 KB) Additional Information: full citation, references, citings, index terms

15 Paper session IR-10 (information retrieval): query expansion: Query expansion using term relationships in language models for information retrieval

Jing Bai, Dawei Song, Peter Bruza, Jian-Yun Nie, Guihong Cao

October 2005 **Proceedings of the 14th ACM international conference on Information and knowledge management CIKM '05**

Publisher: ACM Press

Full text available: pdf(428.57 KB) Additional Information: full citation, abstract, references, index terms

Language Modeling (LM) has been successfully applied to Information Retrieval (IR). However, most of the existing LM approaches only rely on term occurrences in documents, queries and document collections. In traditional unigram based models, terms (or words) are usually considered to be independent. In some recent studies, dependence models have been proposed to incorporate term relationships into LM, so that links can be created between words in the same sentence, and term relationships (e.g. ...

Keywords: information flow, language model, query expansion, term relationships

16 Term clustering of syntactic phrases

D. D. Lewis, W. B. Croft

December 1989 **Proceedings of the 13th annual international ACM SIGIR conference on Research and development in information retrieval**

Publisher: ACM Press

Full text available: pdf(1.62 MB) Additional Information: full citation, abstract, references, citings, index terms

Term clustering and syntactic phrase formation are methods for transforming natural language text. Both have had only mixed success as strategies for improving the quality of text representations for document retrieval. Since the strengths of these methods are complementary, we have explored combining them to produce superior representations. In this paper we discuss our implementation of a syntactic phrase generator, as well as our preliminary experiments with producing phrase clusters. Th ...

17 TextTiling: segmenting text into multi-paragraph subtopic passages

Marti A. Hearst

March 1997 **Computational Linguistics**, Volume 23 Issue 1

Publisher: MIT Press

Full text available: pdf(2.46 MB) Additional Information: full citation, abstract, references, citings

Publisher Site

TextTiling is a technique for subdividing texts into multi-paragraph units that represent passages, or subtopics. The discourse cues for identifying major subtopic shifts are patterns of lexical co-occurrence and distribution. The algorithm is fully implemented and is shown to produce segmentation that corresponds well to human judgments of the subtopic boundaries of 12 texts. Multi-paragraph subtopic segmentation should be useful for many text analysis tasks, including information retrieval and ...

18 Text Categorization: Topic difference factor extraction between two document sets and its application to text categorization

Takahiko Kawatani

August 2002 **Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval**

Publisher: ACM Press

Full text available: pdf(249.49 KB) Additional Information: full citation, abstract, references, index terms

To improve performance in text categorization, it is important to extract distinctive features for each class. This paper proposes topic difference factor analysis (TDFA) as a method to extract projection axes that reflect topic differences between two document sets. Suppose all sentence vectors that compose each document are projected onto

projection axes. TDFA obtains the axes that maximize the ratio between the document sets as to the sum of squared projections by solving a generalized eigenv ...

19 Fast and effective text mining using linear-time document clustering





Bjornar Larsen, Chinatsu Aone
August 1999 **Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining**
Publisher: ACM Press
Full text available: pdf(897.82 KB) Additional Information: full citation, references, citings, index terms

Keywords: clustering, multi-document summarization, text mining

20 Building efficient and effective metasearch engines





Weiyi Meng, Clement Yu, King-Lup Liu
March 2002 **ACM Computing Surveys (CSUR)**, Volume 34 Issue 1
Publisher: ACM Press
Full text available: pdf(416.07 KB) Additional Information: full citation, abstract, references, citings, index terms

Frequently a user's information needs are stored in the databases of multiple search engines. It is inconvenient and inefficient for an ordinary user to invoke multiple search engines and identify useful documents from the returned results. To support unified access to multiple search engines, a metasearch engine can be constructed. When a metasearch engine receives a query from a user, it invokes the underlying search engines to retrieve useful information for the user. Metasearch engines have ...

Keywords: Collection fusion, distributed collection, distributed information retrieval, information resource discovery, metasearch

Results 1 - 20 of 200 Result page: **1** 2 3 4 5 6 7 8 9 10 next

The ACM Portal is published by the Association for Computing Machinery. Copyright © 2006 ACM, Inc.
Terms of Usage Privacy Policy Code of Ethics Contact Us

Useful downloads: Adobe Acrobat QuickTime Windows Media Player Real Player